SOCIAL MEDIA, AI SURVEILLANCE, AND HUMAN DIGNITY: A POST-HUMANIST SYNTHESIS

Joshy N George
Department of Visual Communication, Faculty of Science and Humanities, SRM Institute of Science and Technology, Kattankulathur, TN, India.

Dr Rajesh R
Department of Visual Communication, Faculty of Science and Humanities, SRM Institute of Science and Technology, Kattankulathur, TN, India.

DOI: https://doi.org/10.34293/shanlax.9789361632587.ch031

Abstract

The paper creates a post-humanist framework for social media that demonstrates how people and algorithms and data work together to form communication systems and shows that human dignity emerges through technical and organizational design decisions. The platform surveillance system functions as an analytical sequence which includes data collection and representation followed by forensic analysis and ends with ranking and feedback generation. The analysis reveals four persistent dignity risks which include (1) instrumentalization and manipulation and (2) profiling and automated judgments that create stigma and (3) violations of contextual integrity and (4) expression suppression through chilling effects. The research provides evidence-based solutions to address these negative effects through dignity-centred contracting, redress, dignity-first design principles, dignity impact assessments and bounded inference in governance and calibrated transparency and oversight systems. The resulting framework protects people through its operational structure which prevents them from becoming predictable profiles.

Keywords: *posthumanism, human dignity, AI surveillance, social media*

Introduction

Social media increasingly depends on artificial intelligence (AI) to sense, infer, and steer user activity. The systems present an essential moral conflict regarding the proper method for platforms to deliver customized services while protecting users and preserving human dignity. The post-humanist perspective brings value to the table because it shows how media functions as a collaborative platform that enables people to work with data and machines to generate results. The environment defines dignity as a system condition that emerges from technical and organizational decisions which affect human interactions with platforms.

The current discussions about dignity exist in multiple opposing directions. The two theories of posthumanism and religious humanism establish contrasting views about dignity, yet they enable productive discussions between them (Calo, 2023). The concept of posthumanism does not eliminate human accountability because it shows how power operates between humans and nonhumans in AI systems which requires new perspectives on human and nonhuman relationships (Mellamphy, 2021). The authors support dignity as a universal moral standard which should lead AI development and management because it safeguards human self-governance and logical choice-making (Teo, 2023; Ulgen, 2022).

The posthuman data subject requires legal and policy recognition because their identity emerges from the combination of profiles and traces and algorithmic inferences which challenge established definitions of personhood and online data erasure (Käll, 2017).

The chapter unites these components to demonstrate (a) AI surveillance operations on social platforms (b) the dignity-related risks which result from these operations and (c) practical design and governance solutions derived from the cited research.

1. Conceptual Frame: Posthumanism and Dignity

According to Mellamphy (2021) posthumanism begins with the recognition that human beings develop through technological relationships because their existence arises from these systems rather than preceding them. The new understanding allows us to identify dignity violations through alternative methods. The process of identity creation between data and models makes people lose their dignity when treated as prediction targets for control purposes even when no individual actor means to cause harm (Calo, 2023; Käll, 2017). AI practice benefits from dignity as a practical north star because it protects autonomy and self-presentation integrity and vulnerability (Teo, 2023; Ulgen, 2022).

2. How AI Surveillance Works on Social Platforms

The current implementations differ from one another, yet researchers have developed functional approaches for this technology.

2.1 Capture and Representation.

Machine-readable features and embeddings emerge from the combination of behavioural streams (clicks, dwell time) and media inputs and device signals. AI systems detect synthetic content including deepfakes to maintain information integrity. (Kingsley et al., 2024; Ara et al. 2024).

2.2 Inference and Forensics.

Natural language processing techniques enable the detection of topics and sentiment and risk and deceptive or extremist patterns which support both ranking and safety workflows (Bokolo & Liu, 2024). Object detection and tracking systems operate together to perform video-based surveillance and large-scale scene understanding. (Litoriya et al. 2022)

2.3 Ranking/Targeting and Feedback.

The system depends on inferred signals to generate recommendations for both recommenders and ad systems that optimize video viewing duration and conversion success. User exposure to content leads to a perpetual feedback process which affects their upcoming behaviour. AI technology for user support and moderation functions has become essential for platform operations because platforms need these systems to operate (Alawneh et al., 2023).

The pipeline demonstrates the surveillance process which uses an inference-based orchestration system to perform continuous sensing and classification and ranking operations for user visibility and user perception establishment.

3. Dignity Risks in AI-Mediated Social Media

Four overlapping risk families recur in the literature:

3.1 Instrumentalization and Manipulation.

The implementation of optimization logics which treats users as revenue sources or engagement mechanisms results in a violation of their dignity. The use of manipulative nudges and covert steering methods works against both autonomy and rational agency according to Ienca (2023), Ulgen (2022), Rallabandi et al. (2023) and Kavoliunaite-Ragauskiene (2024).

3.2 Profiling, Stigma, and Automated Judgments.

The classification systems produce essential data which results in permanent factual records about reputation. The systems create more discrimination which strengthens stereotypes while blocking access to services thus violating dignity and equality principles (Orwat, 2024; Silveira & da Silva, 2024).

3.3 Contextual Integrity Violations.

Social limit breakdown occurs when signals from one environment are used for purposes beyond their original intent (such as using social network data for advertising). The absence of context information leads to difficulties for users to handle their digital presence which negatively impacts their basic dignity according to Ienca (2023) and Teo (2023).

3.4 Chilling Effects and Self-Censorship.

Perceived surveillance and the risk of misclassification can suppress expression, particularly among vulnerable groups, narrowing participation and harming well-being (Ulgen, 2022; Silveira & da Silva, 2024).

System design elements such as objectives and datasets and thresholds and interfaces create these harms according to post-humanist analysis (Calo, 2023; Mellamphy, 2021).

4. Design and Governance for Dignity: Actionable Measures

The sources show how technical systems maintain dignity through operational mechanisms during their operational activities.

4.1 Dignity-centred Contracting and Redress.

The readability of contract and policy instruments needs to be combined with fair complaint and appeal procedures that help vulnerable users (de Rooy, 2024).

4.2 Dignity-First Design Principles.

The system requires development to establish spaces where humans can experience vulnerability and share responsibility and gain empowerment through their normal AI interactions (Fasoro, 2024).

4.3 Governance Frameworks with Dignity Checks.

The implementation of Dignity Impact Assessments needs to happen alongside transparency measures and sensitive inference restrictions and specific optimization objective boundaries (Kiden et al. 2024; Kottur, 2024; Saxena, 2024).

4.4 Transparency and Accountability

The research examines two vital areas which focus on Transparency and Accountability in High-Stakes Settings.

The lessons learned about oversight and disclosure and risk management in healthcare settings apply to other fields because unregulated testing breaks trust and damages human dignity (Comeau et al., 2025).

4.5 Information Autonomy.

The ability to control information flow becomes possible through specific tools and practices which Chen (2024) states help people fight extractive dynamics.

The integrated system design evolves from damage response into an active framework for maintaining human dignity in human-machine systems.

5. Brief Vignettes (Applied Illustrations)

5.1 Teen Feeds and Well-Being.

Debates that reduce youth mental health to simple "screen time" narratives are misleading; platform design and recommender choices matter and should reflect adolescents' perspectives and values (Sewall & Parry, 2024; Flitsch, 2022; Figueroa et al., 2024).The system would stop dangerous use patterns by using exposure diversity and friction mechanisms which would protect user autonomy and connection as described by Flitsch (2022) and Figueroa et al. (2022).

5.2 From Risk to Resource.

Digital media supports dignity goals through its orientation toward support and destigmatization when governance implements educational programs and protective measures and enables user agency (Hamilton et al. 2024).

Risk does not determine destiny because design choices and governance systems have the power to transform final results according to the vignettes.

6. Objections and Replies

Objective 1:

The first objection states that surveillance serves two essential purposes which are to ensure safety and maintain quality standards.

Reply:

The practice of concealing information and maintaining control over people leads to distrust between individuals although safety and quality goals remain legitimate. The system requires specific privacy protection measures and disclosure practices to maintain its legitimacy (Lindorfer, 2023).

Objective 2:

The second argument claims users want personalized experiences but transparency would create problems for their overall user experience.

Reply:

Research indicates that personalized services succeed in their objectives yet people experience discomfort because these methods need to gather large amounts of personal information (Lee et al., 2024). The task is calibrated transparency (specific, digestible disclosures) paired with controls that fit users' goals (Schelenz et al., 2020).

Objective 3:

The trade between privacy and personalization is unavoidable according to this argument.

Reply:

The study shows that users experience a sense of conflict between the two systems and information overload creates difficulties for users to understand the content (Garcia-Rivadulla, 2016). The design-for-dignity strategies of bounded inference and consent to purpose and value-sensitive defaults enable the trade-off reduction while keeping user autonomy at the core of the system (Schelenz et al., 2020; Teo, 2023).

Conclusion

Social platforms use AI surveillance to perform an inference-based process which includes data collection followed by sorting and prioritization. The systems which link human beings to these systems decide whether dignity stays whole or gets destroyed based on post-humanist analysis. The global standard of dignity protects autonomy and respect in human relationships, but people now exist as data subjects whose personal information resides in digital records (Calo, 2023; Kall, 2017; Ulgen, 2022). The implementation of dignity-aware platforms requires four essential elements: (1) restrictions on the use of sensitive information and manipulation techniques (2) systems that provide specific information about their operations and methods for users to seek remedies (3) governance frameworks that include dignity evaluation mechanisms and performance boundaries (4) value-based recommender systems that focus on user well-being (de Rooy, 2024; Fasoro, 2024; Kiden et al., 2024; Flitsch, 2022; Figueroa et al., 2024).The system design enables human decision-making to operate alongside machine assistance while maintaining full human control.

References

1. Alawneh, Y. J., Al-Momani, T., Salman, F. N., Al-Ahmad, S. D., Kaddumi, T. A., & Al-Dlalah, M. (2023). A detailed study analysis of artificial intelligence implementation in social media applications (pp. 1191–1194). https://doi.org/10.1109/icacite57410.2023.10182840
2. Ara, A., Alam, M. S., Kamrujjaman, & Mifa, A. F. (2024). A comparative review of AI-generated image detection across social media platforms. https://doi.org/10.62304/jieet.v3i01.77
3. Bokolo, B., & Liu, Q. (2024). Artificial intelligence in social media forensics: A comprehensive survey and analysis. *Electronics*. https://doi.org/10.3390/electronics13091671
4. Calo, Z. R. (2023). Human dignity after the human (pp. 257–268). Springer International Publishing. https://doi.org/10.1007/978-3-031-42127-3_18
5. Chen, D. (2024). The critical canvas—How to regain information autonomy in the AI era. https://doi.org/10.48550/arXiv.2411.16193

6. Comeau, D. S., Bitterman, D. S., & Gallifant, J. (2025). Preventing unrestricted and unmonitored AI experimentation in healthcare through transparency and accountability. *NPJ Digital Medicine, 8*(1). https://doi.org/10.1038/s41746-025-01443-2
7. de Rooy, R. (2024). Contracts for dignity (pp. 74–89). https://doi.org/10.1017/9781009437042.009
8. Fasoro, A. (2024). Cultivating dignity in intelligent systems. *Philosophies, 9*(2), 46. https://doi.org/10.3390/philosophies9020046
9. Figueroa, C. A., Marin, L., de Reuver, M., & Jaff, M. (2024). Translating well-being values into design features of social media platforms: A value sensitive design approach. https://doi.org/10.31219/osf.io/wxtj7
10. Flitsch, S. L. (2022). Inclusive ethical design for recommender systems. https://doi.org/10.48550/arXiv.2209.13021
11. Garcia-Rivadulla, S. (2016). Personalization vs. privacy: An inevitable trade-off? *IFLA Journal, 42*(3), 227–238. https://doi.org/10.1177/0340035216662890
12. Hamilton, J. L., Torous, J., Szlyk, H. S., Biernesser, C., Kruzan, K. P., Jensen, M., Reyes-Portillo, J., Primack, B. A., Zelazny, J., & Weigle, P. (2024). Leveraging digital media to promote youth mental health: Flipping the script on social media-related risk. *Current Treatment Options in Psychiatry*. https://doi.org/10.1007/s40501-024-00315-y
13. Ienca, M. (2023). On artificial intelligence and manipulation. *Topoi, 42*, 833–842. https://doi.org/10.1007/s11245-023-09940-3
14. Käll, J. (2017). A posthuman data subject? The right to be forgotten and beyond. *German Law Journal, 18*(5), 1145–1162. https://doi.org/10.1017/S2071832200022288
15. Kavoliūnaitė-Ragauskienė, E. (2024). Artificial intelligence in manipulation: The significance and strategies for prevention. *Baltic Journal of Law & Politics, 17*(2), 116–141. https://doi.org/10.2478/bjlp-2024-00018
16. Kiden, S., Stahl, B. C., Townsend, B., Maple, C., Charles, V., Sampson, F. A., Gilbert, G., Smith, H., Deshmukh, J., Ross, J., Williams, J., Martínez del Rincón, J., Lisinska, J., O'Shea, K. S., Da Costa Abreu, M., Bencomo, N., Deb, O., Winter, P., Li, P., … Yazdanpanah, V. (2024). Responsible AI governance: A response to UN interim report on governing AI for humanity. https://doi.org/10.48550/arXiv.2412.12108
17. Kingsley, S., S P, A., Babu, B., & Vaithilingam A, H. (2024). AI simulated media detection for social media. https://doi.org/10.47392/irjaeh.2024.0131
18. Kottur, R. (2024). Responsible AI development: A comprehensive framework for ethical implementation in contemporary technological systems. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 10*(6), 1553–1561. https://doi.org/10.32628/cseit241061197
19. Lee, S., Moon, W.-K., & Song, Y. G. (2024). Is it transparent or surveillant? The effects of personalized advertising and privacy policy on advertising effectiveness. *Journal of Interactive Advertising*, 1–21. https://doi.org/10.1080/15252019.2024.2417422

20. Lindorfer, M. (2023). The threat of surveillance and the need for privacy protections (pp. 593–609). https://doi.org/10.1007/978-3-031-45304-5_37
21. Litoriya, R., Ramchandani, D., Moyal, D., & Bothra, D. (2022). Integrated and deep learning–based social surveillance system: A novel approach. *Journal of Automation, Mobile Robotics & Intelligent Systems, 16*, 30–39. https://doi.org/10.14313/jamris/3-2022/22
22. Mellamphy, N. B. (2021). Humans "in the loop"?: Human-centrism, posthumanism, and AI. *Nature and Culture, 16*(1), 11–27. https://doi.org/10.3167/nc.2020.160102
23. Orwat, C. (2024). Algorithmic discrimination from the perspective of human dignity. *Social Inclusion*. https://doi.org/10.17645/si.7160
24. Rallabandi, S., Kakodkar, I. G. S., & Avuku, O. (2023). Ethical use of AI in social media (pp. 1–9). https://doi.org/10.1109/iwis58789.2023.10284706
25. Saxena, A. (2024). AI in governance and policy making. *International Journal of Science and Research, 13*(5), 1218–1223. https://doi.org/10.21275/sr24519015426
26. Schelenz, L., Segal, A., & Gal, K. (2020). Best practices for transparency in machine-generated personalization. In *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 23–28). https://doi.org/10.1145/3386392.3397593
27. Sewall, C., & Parry, D. A. (2024). Social media and youth mental health: Simple narratives produce biased interpretations. *Journal of Psychopathology and Clinical Science*. https://doi.org/10.1037/abn0000950
28. Teo, S. (2023). Human dignity and AI: Mapping the contours and utility of human dignity in addressing challenges presented by AI. *Law, Innovation and Technology, 15*, 241–279. https://doi.org/10.1080/17579961.2023.2184132
29. Ulgen, O. (2022). AI and the crisis of the self (pp. 9–33). Routledge eBooks. https://doi.org/10.4324/9781003030928-3